* BOZEMAN

RADC-TR-60-71 CONTRACT AF 30(602)1818

A Comparison of Five Articulation Tests

by

NICKERSON, MILLER AND SHYNE

MARCH 1960

Prepared for

ROME AIR DEVELOPMENT CENTER
AIR RESEARCH AND DEVELOPMENT COMMAND
U. S. Air Force
Griffiss Air Force Base
NEW YORK

electronics research laboratory

ENDOWMENT AND RESEARCH FOUNDATION AT
MONTANA STATE COLLEGE, BOZEMAN, MONT.

RADC-TR-60-71

FINAL REPORT

A COMPARISON OF FIVE ARTICULATION TESTS

Prepared By

James F. Nickerson, Consultant, Adam W. Miller, Jr. Consultant and N. A. Shyne, Project Leader, (with the cooperation of Dr. William E. Montague, Dr. John Webster and Roy G. Klumpp, U. S. Navy Electronics Laboratory, San Diego, California.)

ELECTRONICS RESEARCH LABORATORY
MONTANA STATE COLLEGE
Bozeman, Montana

March, 1960

CONTRACT AF 30(602)-1818

Prepared For

Rome Air Development Command

Air Research and Development Command

United States Air Force

Griffiss Air Force Base

New York

ABSTRACT

It is difficult to make direct comparisons between estimates of intelligibility afforded by experimental studies employing different articulation tests. Articulation testing is at best rather slow and tedious for the experimenter and the subject. This study sought a comparison of intelligibility estimates of five different articulation tests which have been used in a number of research laboratories. From the performance on these five tests a better comparison of intelligibility level from test to test is available. The Fairbanks Rhyme test as recorded for this study (50 words, administered at the rate of one word every 2½ seconds, for a total of 125 seconds per test) is shown to yield a representative measure of intelligibility with half or less of the experimental time required by the other tests employed. It will therefore be adopted for the articulation testing of Phase II of this study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

ERL/MSC

A COMPARISON OF FIVE ARTICULATION TESTS[1]

I. INTRODUCTION

The development of articulation tests and the refinement of estimates of intelligibility in voice communications has brought marked improvement in the evaluation of communication systems and components. However, two shortcomings are apparent: the lack of adequate comparability from test to test, and the tedium of the experimental task required of the subject. The former makes it difficult to make comparisons of study data from different laboratories and experimenters as to absolute levels of intelligibility. The latter poses a real problem to the experimenter in maintaining the motivation of his subjects and obtaining the maximum valid information per test session.

Further study is needed to identify those test materials which provide adequate estimates of intelligibility, yet which are parsimonious of time and energy of the subject. A better means of transformation of intelligibility levels from test to test and laboratory to laboratory is also needed.

---

[1] Undertaken with the cooperation of Dr. William Montague, Dr. John Webster and Roy G. Klumpp of the Human Factors Division, U. S. Navy Electronics Laboratory, San Diego.

II. DISCUSSION: THE STUDY

The purpose of Phase One of the present study was to compare several standard articulation tests employed in recent intelligibility studies carried on in several of the leading research laboratories. From these data it is hoped: 1) to establish a means for equating intelligibility levels from experiment to experiment or from laboratory to laboratory; 2) to select from among the articulation tests employed, a single test which will yield representative intelligibility estimates with the greatest efficiency of experimental time and effort.

1. EQUIPMENT

The design of the articulation testing facility used in this study is shown in Figure 1. A detailed description of the test facility is to be found in a report submitted July 13, 1958, to Rome Air Development Center, Contract AF 30(602)-1818, by Leslie E. McCoy, entitled, "The Articulation Test Facility at Montana State College". A number of new filters and attenuation networks have been added since publication of the above report.

2. EXPERIMENTAL DESIGN

Ten subjects were given five standard articulation tests covering six signal-to-noise ratios* (30 test conditions). The presentation was replicated two times to afford estimates of additional practice effect and

---

*The term "signal-to-noise ratio" here might also be construed to mean "relative noise level", since there is no absolute basis of signal amplitude. However, in this connection the term "signal-to-noise ratio" will be utilized throughout this report. A "signal-to-noise ratio" of zero db was achieved by equating the rms amplitude of the noise with the rms amplitude of the calibration tone at the beginning of each tape. The calibration tone in turn bore a definite relationship to the test material (i.e., the rms amplitude of the calibration tone was made equal to the syllabic peak of the long sound in the word "write", which preceded each test word).

to give greater stability of articulation scores or intelligibility-level estimates. Responses were translated into percent intelligibility and the data analyzed by graphical presentation and analysis of variance to determine the variability among:

1) subjects
2) articulation tests
3) signal-to-noise ratios
4) replications
5) interactions among these factors.[1]

a. Subjects

The ten subjects, six male and four female, ranged in age from 16 to 23. Selection was made following not less than six hours practice, given in two or more test sessions, during which each subject showed acceptable ability to "hear" the test materials under a wide range of noise conditions.

b. Test Materials[2]

The five standard articulation tests were selected in consultation with Dr. William Montague and Dr. John Webster of the Human Factors Division of the U. S. Navy Electronics Laboratory (NEL) in San Diego. All test materials were recorded at NEL under close laboratory supervision,[3] with the exception of the W-22 word lists which were re-recorded from disks available commercially from Central Institute for the Deaf (CID) of St. Louis. The test materials included[2]:

1) Harvard Sentences (HS)
2) Phonetically Balanced Word Lists (PB)
3) Navy Communication Words (NCW)

---

[1] The analysis of variance did not include comparison of replications due to computer program limitations. Comparison of performance by replication is therefore possible only by examination of the graphical presentation at the time of this report. (See Figures 10 through 12).

[2] Sample Word Lists of each test are given in Appendix IV.

[3] All NEL materials were clipped 3 db.

4) Fairbanks Rhyme Test (FB)

5) W-22 Word Lists (W-22)

6) Psuedo Navy Sentences (PNS)

(PNS was included among the test materials for the two replications; however, this last test was not included in the graphical and statistical analysis which follows.)

1) Harvard Sentences (HS)[1]

Test materials were developed by the Harvard Psycho-Acoustic Laboratory (PAL) during World War II. Stimulus sentences were drawn from 100 short sentences, each containing four monosyllables and one disyllable. Each test included 20 test sentences (80 monosyllables and 20 disyllables). Test materials were recorded by NEL at a rate permitting 20 seconds (write-down time) between each sentence (approximately 500 sec. for each test form). Only five different test forms were available to this study.

2) Phonetically Balanced Word Lists (PB)[1]

These materials were developed by the Harvard-Psycho-Acoustic Laboratory during World War II. Stimulus words were drawn from a vocabulary of 200 monosyllabic words divided into 50 words per test form. The materials were recorded by NEL from PAL-PB word lists at a rate of one word each 4 seconds (250 seconds per test form). Ten different test forms were used in this study.

3) Navy Communications Words (NCW)

Test materials were developed from standard Navy communication vocabulary[2] by the U. S. Navy Electronics Laboratory in San Diego. Each test form

---

[1] Egan, J. P., "Articulation Testing Methods", Laryngoscope, 1948a, 58, 955-991.

[2] NCW lists were selected from words in the PNS.

presented 50 words (40 monosyllables and 10 disyllables) from a total vocabulary of 400 monosyllables and 100 disyllables. Materials were recorded by NEL at a rate of one word each five seconds (250 seconds per test form). Ten test forms were used in this study.

4) Fairbanks Rhyme Test (FB)[1]

The Rhyme test is a test developed in the Speech Research Laboratory, University of Illinois. It is of the completion type. Stimulus words are drawn from a vocabulary of 250 common monosyllables (50 sets of five rhyming words each). For each test form one word is drawn from each set. These materials were recorded by NEL with a stimulus word presented each 2 seconds (100 seconds per test form). Twelve different test forms were used in this study.

5) W-22 Word Lists[2] (W-22)

These materials were developed by the Central Institute for the Deaf from the PB word lists. Selection was made to obtain a more uniform difficulty than the original PB lists. Stimulus words are drawn from a vocabulary of 200 one-syllable words divided into 50 words per test form. The materials were recorded by CID at the rate of one stimulus word each 5 seconds (250 seconds per test form). Twenty-four different test forms were available in the study.

---

[1] Fairbanks, G., Test of Phonemic Differentiation: The Rhyme Test, J. Acoust. Soc. Amer., 1958, 30, 596---98.

[2] Hirsh, I.J., et.al., Development of materials for speech audiometry. Journal of Speech and Hearing Disorders, 1952, Vol. 17, No. 3, 328-337. Central Institute for the Deaf, (St. Louis, Mo.) Auditory Test W-22; Phonetically-Balanced Word Lists. (ONR-Contract N60nr-272, Project No. NR 142-170, Task Order III).

6) Psuedo Navy Sentences (PNS) (Employed in replications only)

These test materials were developed by the U.S. Navy Electronics Laboratory at San Diego. Stimulus sentences were drawn from standard Navy vocabulary, each sentence involving 4 monosyllables and 1 disyllable. Each test form includes 20 sentences (80 monosyllables and 20 disyllables). Test materials were recorded by NEL at a rate permitting 20 seconds between each sentence (approximately 500 seconds for each test form). Only five test forms were available to this study.

With the exception of the W-22 Word Lists, each of the test materials used was clipped 3-db and pre-recorded in the recording laboratories at NEL. The W-22 test materials were recorded on tape from the disk recordings commerically available from the Central Institute for the Deaf.

Subjects responded to each test by writing each test word or sentence as presented. Subjects were instructed to write what they thought they heard. Correct responses were transformed into a percentage of total responses to obtain the per cent articulation.

c. Signal-to-Noise Ratios

In order to compare the test materials over a broad range of noise and intelligibility conditions, signal-to-noise ratios were chosen to yield intelligibility levels ranging from 20 to 95 percent. This decision was based on the intelligibility levels yielded by the use of the W-22 word lists in an earlier study[1] and the preliminary results yielded by the practice sessions for the ten subjects. Six signal-to-noise ratios were selected, ranging as follows: 0, -8, -12, -16, -20 db, with the signal level remaining constant. Figure 2 indicates that intelligibility levels ranged from 15% or

---

[1] Nickerson, J.F., "A Study of the Effect of Frequency Translational Error on Intelligibility of Speech in the Presence of Noise", Electronics Research Laboratory Technical Report, Montana State College (1958).

less for all tests at S/N = -20 db to 90% or better on 4 of the 5 tests at a S/N radio of 0 db.

d. Experimental Conditions and Limitations

Testing limitations introduced by the 10-bay test facility required all subjects to listen to the same test condition simultaneously. In the presentation of test conditions there were thirty possible pairings of each of the five different speech materials with each of the six signal-to-noise ratios. The two replications of the test conditions called for re-presentation of the same materials though in a different random order.[1]

e. Detail of Signal Recording and Noise Presentation
(See description of individual tests)

The test material used for each test session was re-recorded (Ampex 350) from the master recordings at constant signal level with reference to a calibration tone supplied on each tape or disk. All signals were then presented to the earphones of each subject at 80-db sound level (re.0002 dyne/cm^2). Calibrated random noise (20 kc bandwidth) from a GR 1390-A noise generator was fed to each earphone and varied in level to produce the desired signal-to-noise ratio.

f. Data Presentation and Analysis

Articulation scores were computed as the ratio of the number of correct responses to the total responses. Mean values were computed for the responses of the ten subjects for each of the 30 test conditions for each replication of the experiment and also for the pooled responses of all three replications. These mean values were then plotted by: test material, replication, and

---

[1] It will be noted that a relatively serious limitation was presented by the fact that only 5 different recorded test forms were available for the Harvard Sentences while from 10 to 24 recorded forms were available for each of the other tests. This means that in the case of the Harvard Sentences a single recording could have been presented twice per replication (a possible maximum of six presentations), with attendant increase of the possibility of memorization.

subject. An analysis of variance (factorial design) was then run on the data for the second replication to determine the significance of any observed differences attributable to: test materials, subjects, signal-to-noise ratios, and any interaction among these factors. From the graphical data presentations and the analysis of variance, interpretation was sought concerning the comparability of the five articulation tests, and which test or tests appeared to give a representative estimate of intelligibility with the least experimental time and effort.

III. FINDINGS: INTELLIGIBILITY COMPARISONS

1. GRAPHICAL PRESENTATION

a. By Test

Figures 2 through 10 (Appendix III) present the data of variations among the five test materials over each of the six S/N ratios. Figures 2 and 2a make possible comparison of the tests on the basis of the data of the combined sessions. In the latter case the curves have been smoothed to suggest the more generalized performance on the five tests. Figures 3, 4 and 5 show the data of the initial testing and the two replications, respectively.

Substantial agreement is shown between FB, PB, and NCW in the data of combined sessions (Fig. 2). All three tests yield test scores approximating the average for the five tests used in the study. The agreement is still apparent in the data of each test session, though variations are more pronounced particularly in the first and third test sessions. W-22 yields substantially lower intelligibility scores.[1] These differences are on the order of 20-30% below the average for the 5 tests. In contrast HS yields higher articulation scores (from 5% to 25% higher than average in the upper intelligibility levels). Under unfavorable listening conditions HS yields as much as 30% higher scores. These observed differences disappear when noise conditions become highly unfavorable and intelligibility approaches zero.

b. By Test Material and Replication

Figures 6 through 10 show the variation to be found in averaged performances of subjects on the five tests for each of three test sessions. Greatest consistency is shown by FB, NCW and W-22, although there is considerable variation evident for all tests. It is for this reason that interpretations are more readily made on the data of combined sessions.

---

[1] However, with the same level presentation to listeners, the unclipped W-22 lists have a lower consonant level than the NEL clipped materials, yielding a lower intelligibility score.

Figure 11 shows the data of average intelligibility for all six S/N ratios for each of the five tests, and each of the three test sessions. The relative levels of estimates are immediately apparent with FB, PB and NCW yielding median values just above 60%, with HS and W-22 yielding values of approximately 80-85% and 35-40%, respectively.

c. By Signal-to-Noise Ratio

Figure 12 shows the comparison of intelligibility levels by S/N ratio. The expected rise in intelligibility, as S/N ratio improves, is apparent up to approximately 50%.

2. STATISTICAL ANALYSIS

An analysis of variance (factorial design) was done for the second of the three testing sessions (See Table 4 for Analysis of Variance data).

The results indicate significant variation among subjects, among signal-to-noise ratios, among test materials and for the interaction between signal-to-noise ratios and test materials (See Tables 1-3 and 5-8 for actual mean values).

The test of significance serves only to substantiate that the differences to be observed graphically between subjects, between signal-to-noise ratios, and between test materials are significant differences which could not have occurred by chance.

The significance of difference due to interaction between signal-to-noise ratios and test materials gives further confidence in interpreting from Figure 2 that there is substantial variation in test performance according to signal-to-noise ratio. This is to be anticipated from the fact that variation would tend to decrease as intelligibility approached maximum (95-100%), and similarly variation would fall to zero when intelligibility levels reached zero. Greatest variation is to be expected around the 50% level of intelligibility, a fact borne out again by these data.

IV. SUMMARY & CONCLUSIONS

1. SUMMARY

Phase One of the present study sought to establish means for equating or transforming levels of intelligibility obtained by the use of several different articulation tests and to aid in selecting an articulation test for future work (Phase Two of this study). The selected test should yield representative measures of intelligibility and offer better efficiency of experimental time and effort.

Ten subjects were given five standard articulation tests over signal-to-noise ratios chosen to yield from 20% to 95% intelligibility. The study was given two replications, using different randomized orders of presentation of test conditions, to ensure greater stability and validity of articulation scores. Data were analyzed both graphically and statistically.

The five tests employed were Harvard Sentences (HS), Phonetically Balanced Word Lists (PB), Navy Communication Words (NCW), Fairbanks Rhyme Test (FB), and W-22 Word Lists (W-22). A sixth test, Psuedo Navy Sentence. (PNS), was employed in the second and third test sessions (replications), but has not been analyzed at the time of this report.

The data reveals that FB, PB, and NCW yield similar scores at all levels of intelligibility, which approximate the averaged scores for all five tests employed. HS yields substantially higher scores and W-22 yields substantially lower scores as compared to the FB, PB and NCW.

The materials of the Fairbank Rhyme Test (FB) require one-half of the time required to administer the PB, NCW, or W-22 lists, and approximately one-fourth of the time required for HS or PNS. Yet the FB lists yield representative scores.

2. CONCLUSIONS

Although there is significant variation between subjects and evident variation between the three test sessions, the averaged scores of the three test sessions (3x10 = 30 test scores per test condition) appear to give a stable and valid estimate of intelligibility.

Comparison of levels of intelligibility is possible by means of the smoothed curves which suggest the more generalized performance on each of the five arti .ulation test used.

The Fairbanks Rhyme test closely approximates the average test scores at all levels of intelligibility, and it requires but one-half as much administration time as the next shorter test. On these bases the FB test was accepted as the most useful and efficient test of the five investigated.

The Fairbanks Rhyme test will therefore be used on all subsequent articulation testings on this project.

APPENDICES

APPENDIX I

EQUIPMENT DESIGN

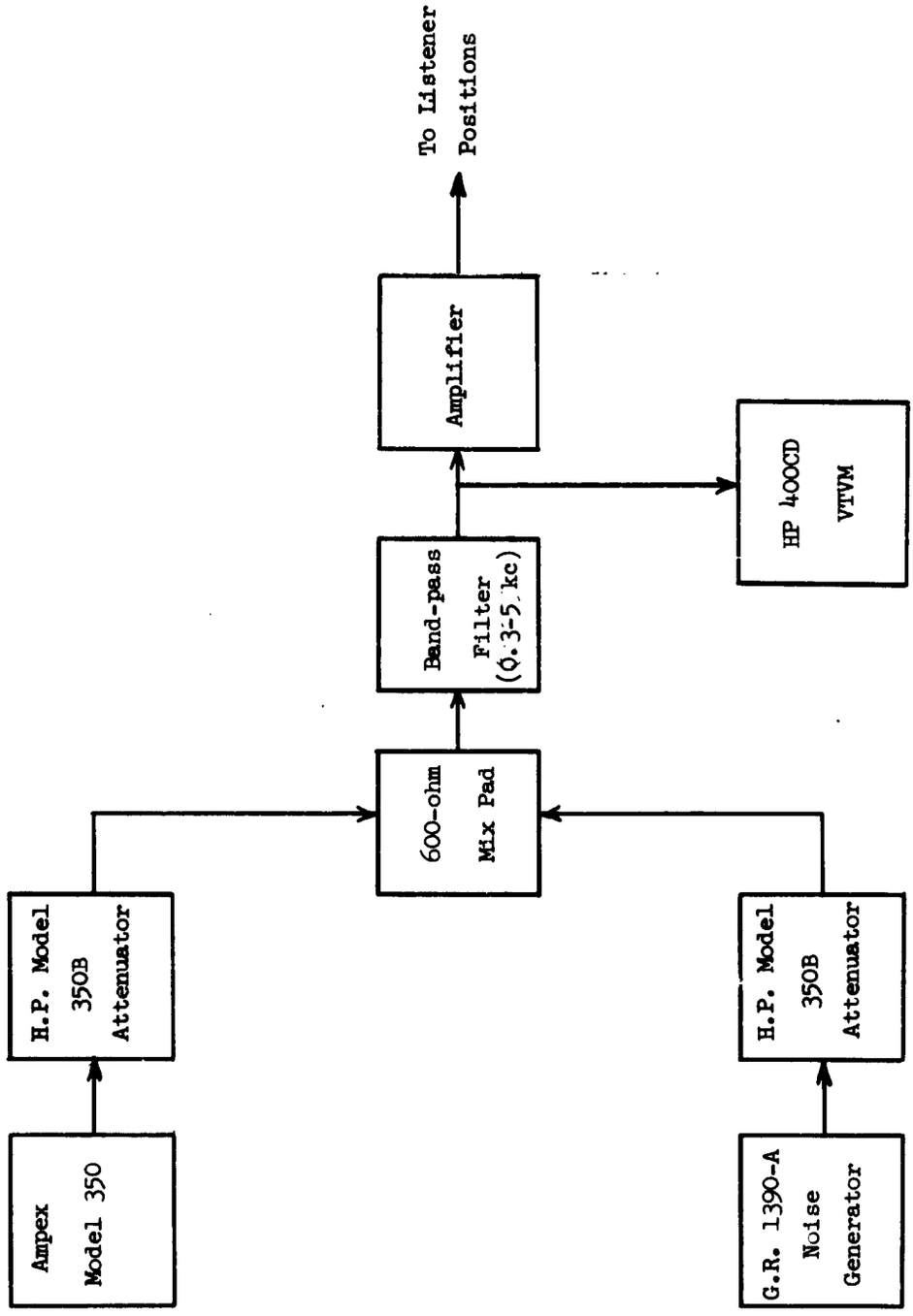

Figure 1. Block Diagram of Equipment for Listener Tests

APPENDIX II

TABLES

TABLE I

TABLE OF MEANS: INITIAL TESTING SESSION BY SUBJECT

	S/N						
Subjects	0	-4	-8	-12	-16	-20	All
1	80.2	78.6	79.0	50.8	32.8	04.4	54.3
2	78.8	79.0	70.2	57.6	28.6	07.6	53.8
3	81.6	83.2	78.2	59.2	34.6	12.0	58.1
4	82.0	84.4	77.0	55.0	32.0	06.4	56.1
5	80.0	81.2	76.0	62.2	29.2	05.6	55.7
6	78.0	85.0	68.6	51.8	28.2	06.4	53.0
7	83.0	82.4	76.4	61.8	32.2	03.2	57.5
8	81.6	76.4	72.4	58.4	29.4	08.0	54.4
9	76.8	80.4	69.2	45.2	25.4	04.4	50.2
10	79.0	79.4	75.6	45.4	30.4	10.0	53.3
All	80.1	81.0	73.3	54.1	30.[illegible]	06.8	54.6

TABLE 2

TABLE OF MEANS: FIRST REPLICATION BY SUBJECT

	S/N						
Subjects	0	-4	-8	-12	-16	-20	All
1	89.6	82.0	83.6	69.2	69.8	8.0	65.5
2	85.0	80.4	74.2	46.4	25.8	3.2	52.5
3	88.2	80.4	77.2	56.8	34.4	7.2	57.3
4	89.2	79.0	75.8	56.4	26.8	6.0	55.5
5	89.2	82.4	77.8	56.2	33.2	6.2	57.5
6	84.4	72.8	72.0	46.0	25.8	2.4	50.6
7	88.8	77.4	77.8	50.6	33.0	2.0	55.0
8	90.0	77.4	77.4	55.6	31.2	5.2	56.1
9	84.8	75.2	73.2	51.2	29.0	2.4	52.6
10	85.4	77.0	75.4	53.0	28.8	4.4	56.0
All	87.5	78.4	76.4	53.6	32.9	4.7	55.8

TABLE 3

TABLE OF MEANS: SECOND REPLICATION BY SUBJECT

Subjects	S/N						
	0	-4	-8	-12	-16	-20	All
1	85.6	86.0	82.4	49.8	29.6	0.0	55.6
2	88.6	82.8	74.0	43.6	21.4	4.6	52.5
3	90.0	84.4	78.2	51.8	30.2	8.6	57.2
4	89.6	83.8	76.6	51.8	26.0	5.6	55.6
5	92.0	83.6	78.0	51.6	31.8	11.2	58.0
6	85.6	79.2	73.8	47.0	22.8	4.4	52.1
7	92.8	82.6	83.8	54.8	29.0	5.6	56.1
8	90.8	83.2	80.4	51.2	29.2	6.0	56.8
9	87.0	82.8	78.4	43.2	28.0	3.8	53.9
10	84.4	83.2	77.8	46.6	30.4	7.8	55.0
All	88.6	83.2	78.3	49.1	27.8	5.8	55.2

TABLE 4

ANALYSIS OF VARIANCE: FIRST REPLICATION

Source of Variation	°of Freedom	Squares	Squares	F
(A) Subjects	9	4,097.35	455.26	.05
(B) Signal-to-Noise Ratio	6	254,002.71	42,333.79	.001
(C) Test Materials	4	72,168.35	18,042.09	.001
Interactions:				
(A) x (B)	36	1,574.72	43.74	N.S.
(A) x (C)	54	3,319.49	61.47	N.S.
(B) x (C)	24	40,513.09	1,688.05	.001
(A) x (B) x (C)	216	7,679.04	35.55	N.S.
Pooled Error	240	48,192.13	200.80	--
Total	351	469,390.2	---	---

*Significant at the .01 level.

TABLE 5

TABLE OF MEANS: INITIAL TESTING SESSION BY TEST MATERIAL

Test Material	S/N					
	0	-4	-8	-12	-16	-20
PB	74.1	74.2	72.4	48.2	07.8	00.0
HS	98.4	98.6	96.3	76.9	49.0	12.6
FB	95.4	91.4	80.4	47.0	50.8	12.6
NCW	93.4	77.8	73.0	65.2	42.8	08.8
W-22	66.2	63.0	44.2	33.4	01.0	00.0
All	80.1	81.0	73.3	54.1	30.3	06.8

TABLE 6

TABLE OF MEANS: FIRST REPLICATION BY TEST MATERIAL

Test Material	S/N					
	0	-4	-8	-12	-16	-20
PB	92.8	88.8	77.4	48.4	10.8	01.4
HS	99.5	99.6	97.6	90.4	91.8	00.5
FB	85.2	78.4	80.4	61.4	22.6	14.8
NCW	91.2	73.0	82.2	52.4	35.2	05.6
W-22	68.6	52.2	44.6	15.6	04.0	01.2
All	87.5	78.4	76.4	53.6	32.9	04.7

TABLE 7

TABLE OF MEANS: SECOND REPLICATION BY TEST MATERIAL

Test Material	S/N					
	0	-4	-8	-12	-16	-20
PB	94.4	92.0	95.0	51.6	23.8	00.4
HS	99.6	99.0	98.5	90.9	73.4	22.8
FB	87.0	84.6	80.2	29.0	41.6	00.0
NCW	95.6	85.2	79.6	57.2	00.0	05.6
W-22	66.6	55.0	38.4	17.0	00.0	00.0
All	88.6	83.3	78.3	49.1	27.8	05.8

TABLE 8

TABLE OF MEANS: COMBINED SESSIONS BY TEST MATERIAL

Test Material	S/N					
	0	-4	-8	-12	-16	-20
PB	78.1	85.0	81.6	48.4	14.1	0.6
HS	99.2	99.1	97.5	86.1	71.4	11.9
FB	89.2	84.8	80.3	45.8	38.3	9.1
NCW	93.4	78.7	78.3	58.3	26.0	6.7
W-22	67.1	56.7	42.4	22.0	1.7	0.4
All	85.4	80.9	76.0	52.5	30.4	5.8

APPENDIX III

GRAPHICAL PRESENTATION

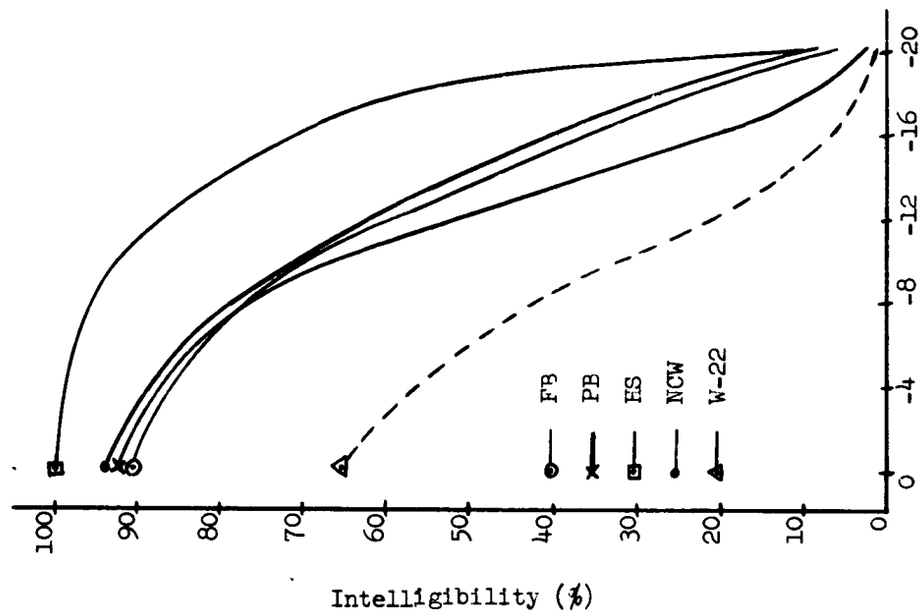

Figure 2a. Intelligibility of Test Materials: Smoothed Curves

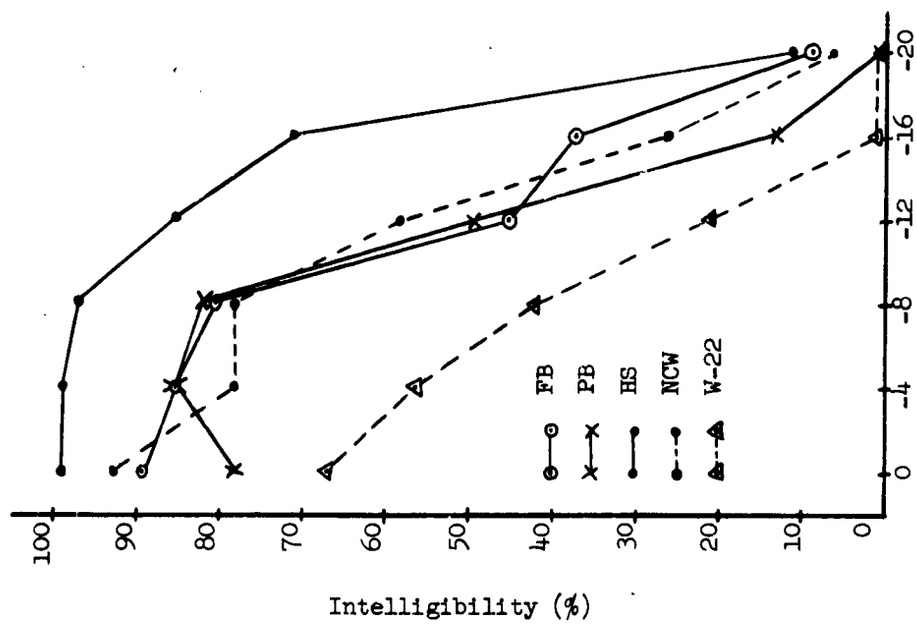

Figure 2. Intelligibility of Test Materials: Combined Sessions

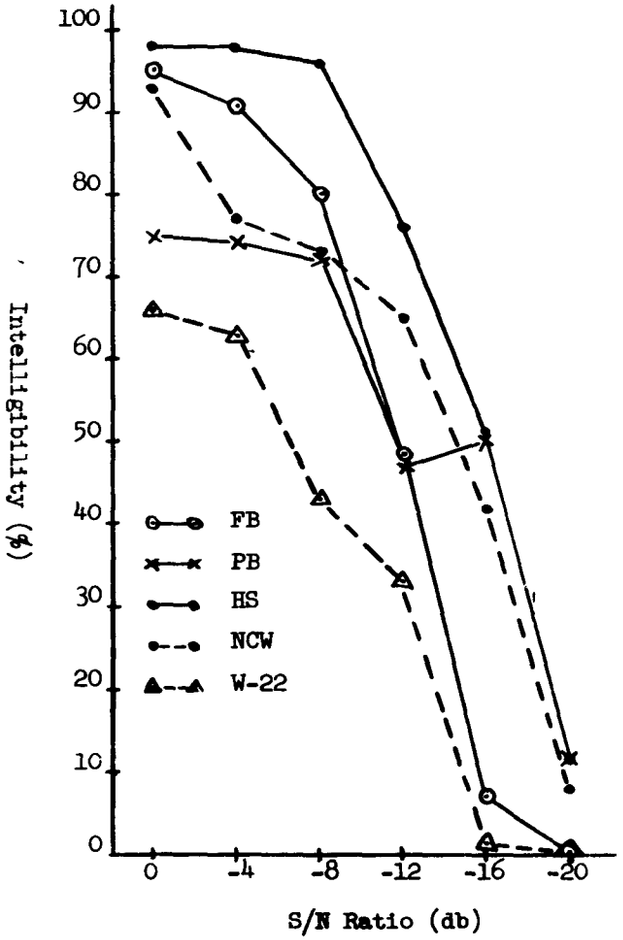

Figure 3. Intelligibility of Test Materials: Initial Testing Session

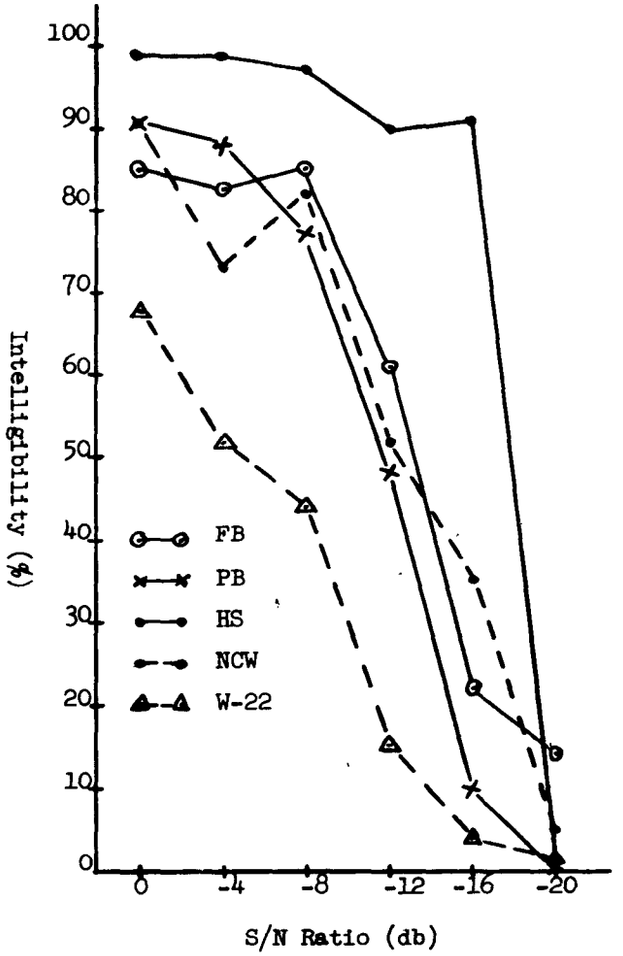

Figure 4. Intelligibility of Test Materials: First Replication

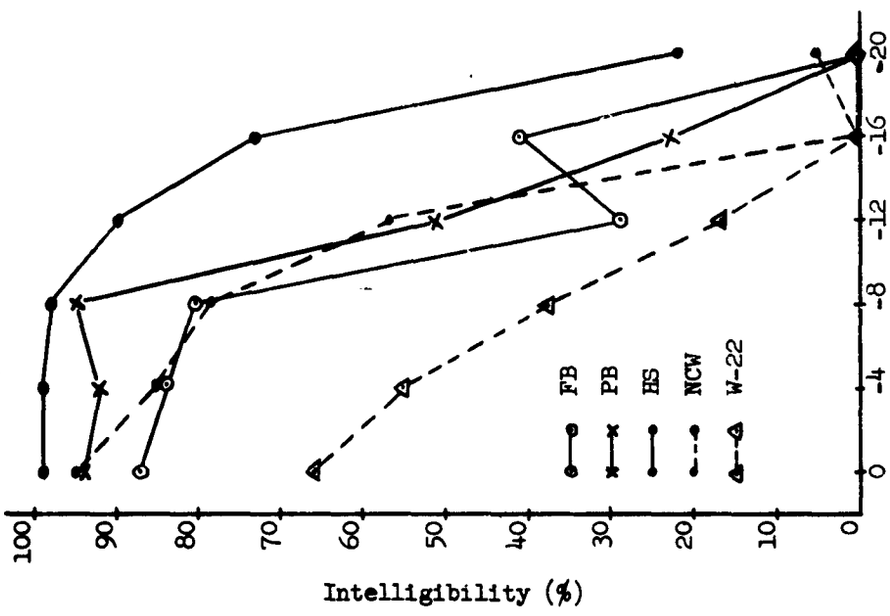

Figure 5. Intelligibility of Test Materials: Second Replication

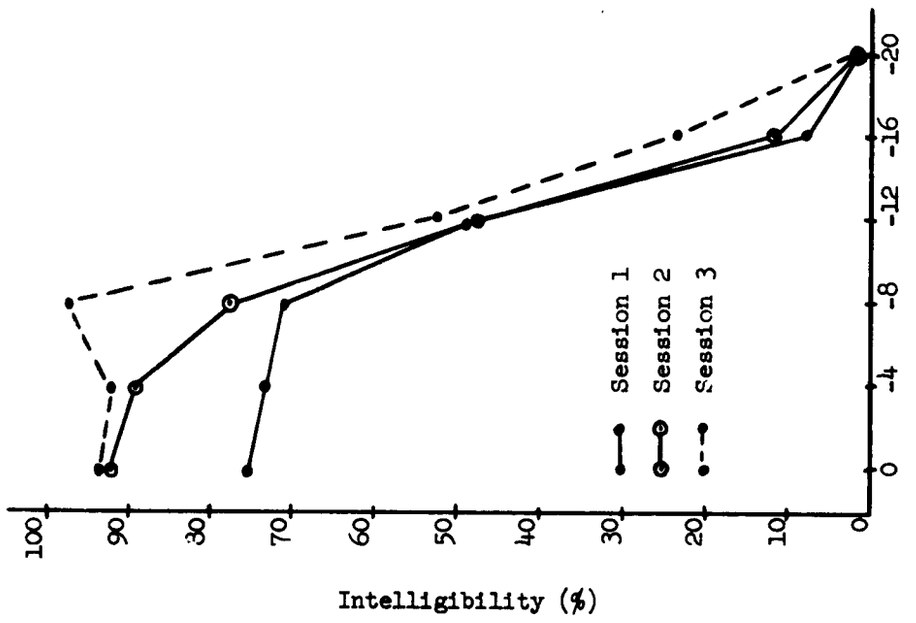

Figure 7. Intelligibility of Phonetically Balanced Word List: By Session

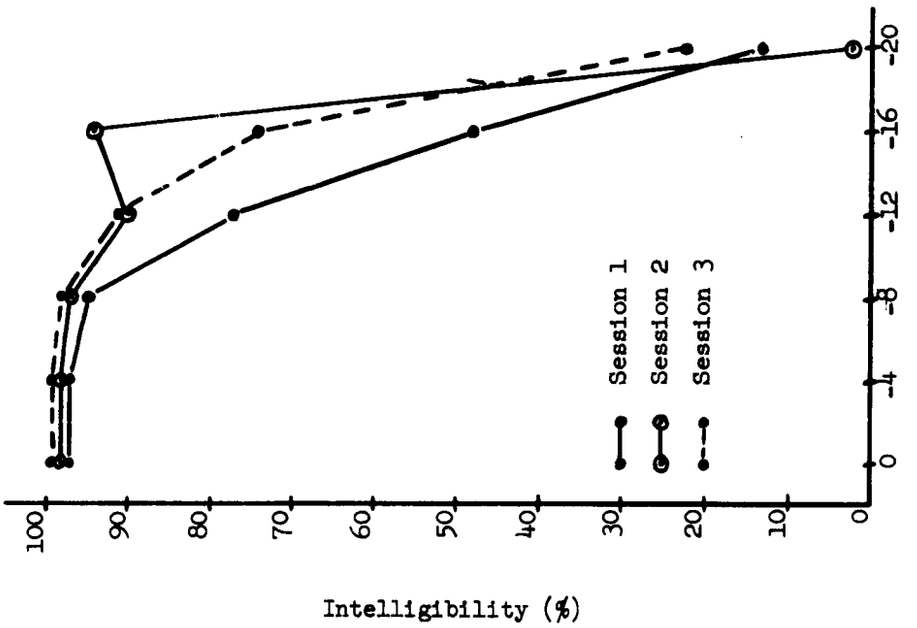

Figure 6. Intelligibility of Harvard Sentences: By Session

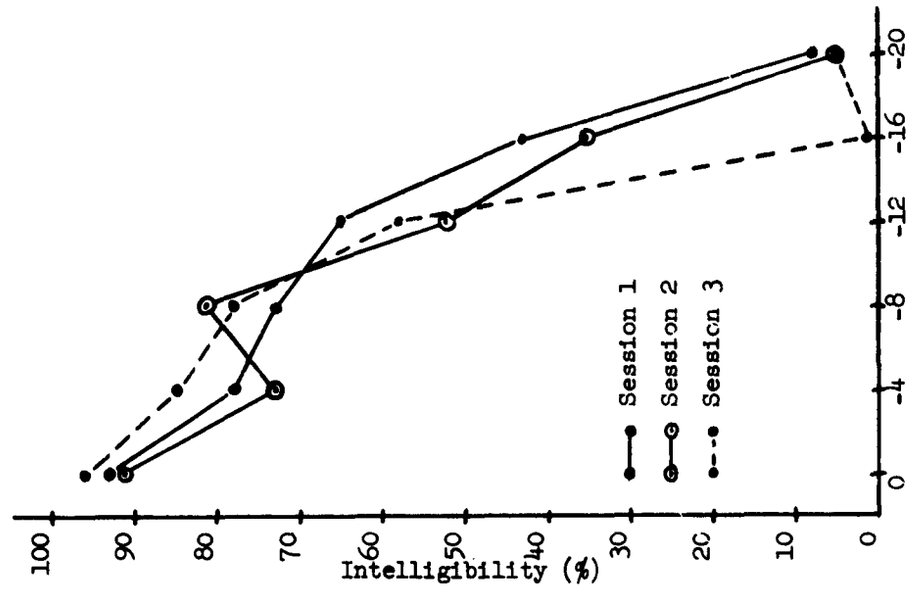

Figure 9. Intelligibility of Navy Communication Words: By Session

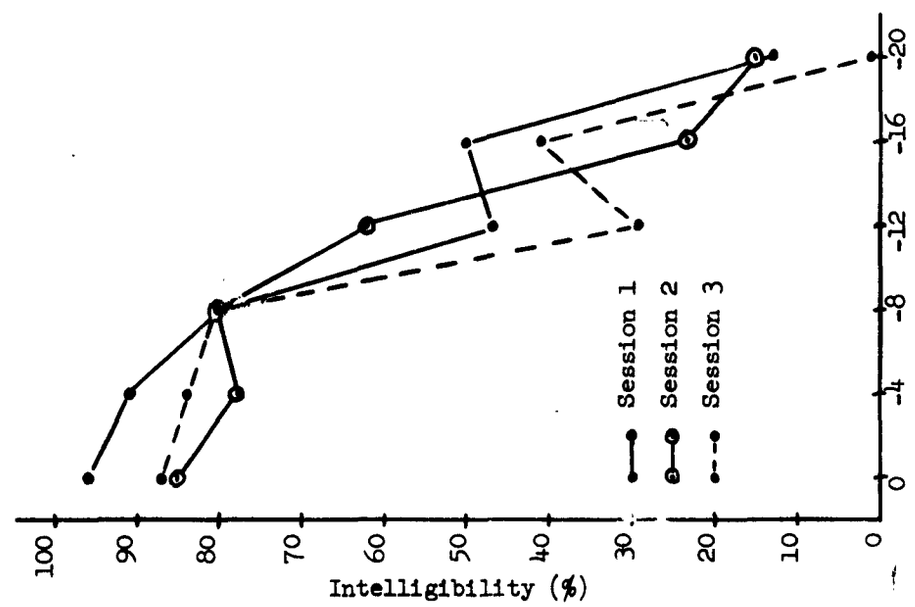

Figure 8. Intelligibility of Fairbanks Rhyme Test: By Session

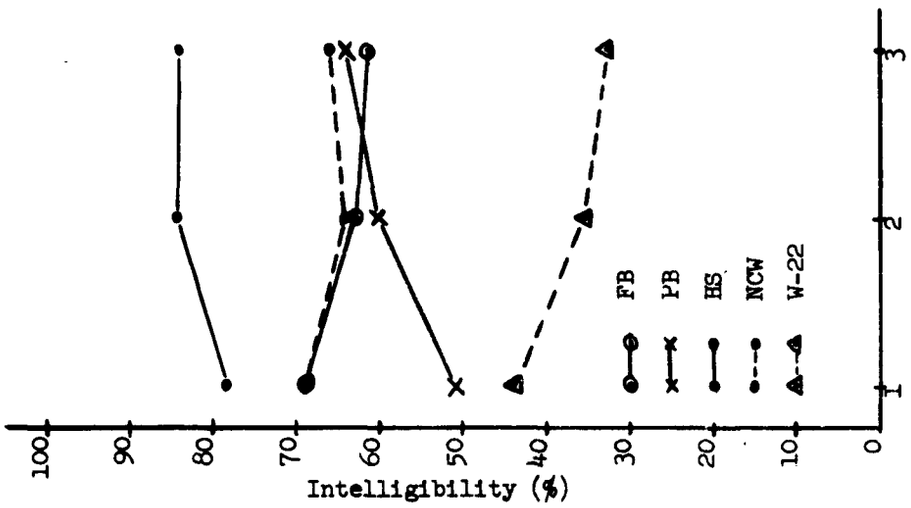

Figure 11. Intelligibility of Test Material: By Session

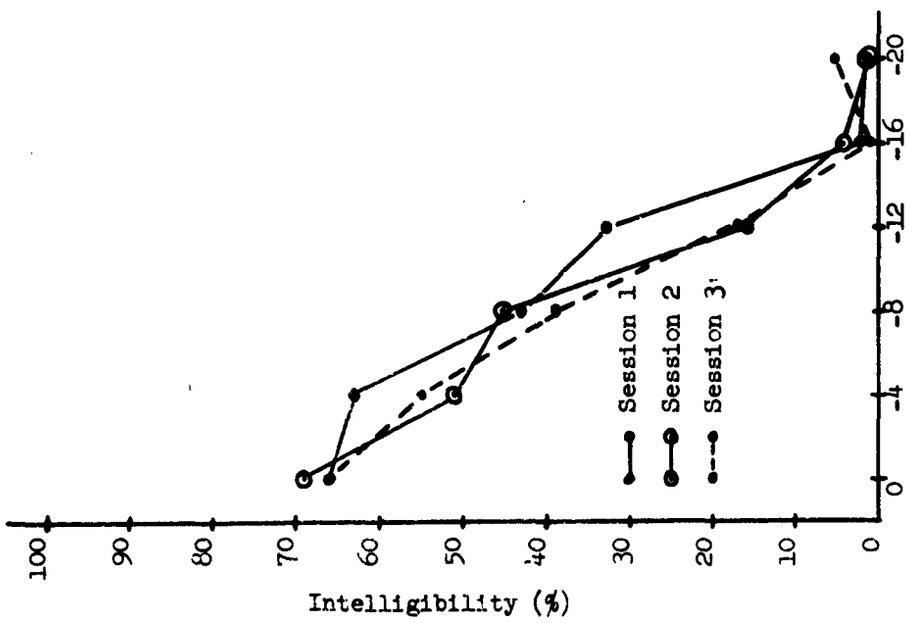

Figure 10. Intelligibility of W-22 Word Lists: By Session

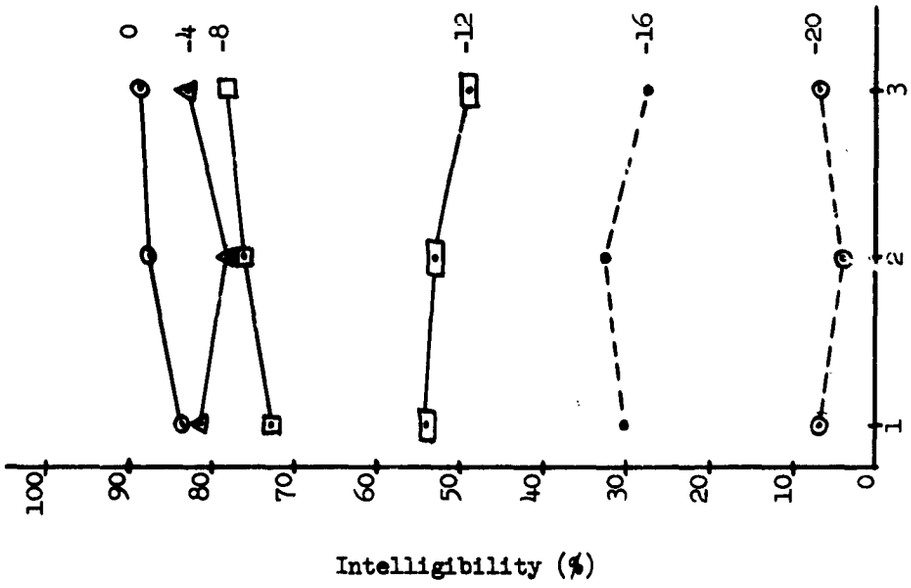

Figure 12. Intelligibility of Test Material By Signal to Noise Ratio and By Session

APPENDIX IV

SAMPLE WORD LISTS

SAMPLE: FAIRBANKS TEST ORDER 4 -- RT-3	SAMPLE: C.I.D. AUDITORY TEST W-22 (PB WORD LISTS) LIST 2A	SAMPLE: WORDS FROM NAVY VOICE COMMUNICATIONS -- LIST 9	SAMPLE: PB 4 ORDER 4
cock	yore (your)	aim	dodge
pack	bin (been)	fight	slap
tin	way (weigh)	whistled	raw
get	chest	east	oils
right	then	give	float
jaw	ease	barked	scab
name	smart	drive	race
sang	gave	still	hatch
sire	pew	downward	shin
day	ice	before	earn
boil	odd	phase	shed
hold	knee	tools	budge
mail	move	ice	peck
lip	now	kick	cloak
run	jaw	hammer	or
lock	one (won)	stole	tick
feel	hit	across	starve
seal	send	gang	pinch
pink	else	white	bath
wine	tare (tear)	slow	blonde
sale	does	float	eel
dig	too (two,to)	run	beast
coon	cap	calm	sketch
gate	with	dye	heed
fast	air (heir)	source	neat
page	and	steam	touch
not	young	being	bus
went	cars	air	rave
beat	tree	droop	fin
born	dumb	frayed	sour
file	that	level	rack
pop	die (dye)	coast	bush
red	show	speed	hiss
lump	hurt	yell	move
tell	own	hulk	test
duck	key	smoke	hot
lark	oak	curse	sage
tore	new (knew)	lead	course
main	live (verb)	passage	new
kid	off	doctor	bee
test	ill	chew	strap
wit	rooms	hop	how
hook	ham	zero	dupe
make	star	less	kite
must	eat	catwalk	frown
will	thin	state	rut
hen	flat	blocked	court
side	well	drill	pert
bend	by (buy)	topside	pod
cod	ail (ale)	count	merge

SAMPLE: [illegible]

LIST B

1. Wrap the body in the red, [illegible]
2. Change the tube and fix the [illegible]
3. Keep the tractor clear of the [illegible]
4. We'll hold jets till we get the [illegible]
5. Be sure to set the chain firmly [illegible]
6. The forward deck [illegible] a top.
7. Drop each plane down the [illegible]
8. Wave your green wand to signal [illegible]
9. The loss of the second fleet [illegible]
10. Seldom have I seen such a fight.
11. Relieve the watch each two hours.
12. We will need to speed up to get the planes off [illegible]
13. He'll pass by once more to windward.
14. Get the red vessel back in line.
15. Send a man forward with torch to cut [illegible]
16. Stand clear of jet intakes and pipes.
17. Get the planes up for [illegible] effort.
18. Send two men to repair the [illegible]
19. Pour sand on that oil to prevent a [illegible]
20. The hook tore off the sight on the [illegible]

SAMPLE: HARVARD LIST 1

1. Watch for small bugs [illegible] soft apples.
2. It is hard to erase blue or red ink.
3. Write at once or you'll sure regret it.
4. The doorknob was made of bright [illegible]
5. The wreck occurred by the [illegible]
6. A pencil with white paint writes well.
7. You might have to coax a calf to drink [illegible]
8. Schools for ladies teach charm and grace.
9. His blank gaze spoke of sheer wonder.
10. The lamp shone with a steady green flame.
11. They took the axe and the saw to the [illegible]
12. Lower the sick boy to the first deck.
13. A storm came with the immense black [illegible]
14. A true stitch saves time and worry.
15. The ancient coin was quite dull and worn.
16. Shout as loud as your lungs allow.
17. The shaky barn fell with a loud crash.
18. Keep the man you have, I advise.
19. Jazz and swing fans like fast music.
20. They opened the desk and saw a big gun.